

SEQUENCE LISTING

DNASIS- DNA Translation [T-INS]
 File Name : T-INS
 Range : 2 - 7286 Mode : Normal
 Codon Table : Universal

5' *Start of
Seq 3 + 4*

GAG	CTG	AGC	TGA	ACT	GGC	CCT	CCT	GGG	GAC	TCA	GCA	AGC	TCT	CTA	GAG	CCC	CCC
E	L	S	*	T	G	P	P	G	D	S	A	S	S	L	E	P	P
ACA	TGC	TCC	CCC	ACC	GGG	GTC	CCC	CGG	TTG	CGT	GAG	GAC	ACC	TCC	TCT	GAG	GGG
T	C	S	P	T	G	V	P	R	L	R	E	D	T	S	S	E	G
CTC	CGC	TCG	CCC	CTC	TTC	GGA	CCC	CGG	GGG	GCC	CCG	GCT	GGC	CAG	AGG	ATG	GAC
L	R	S	P	L	F	G	P	P	G	A	P	A	G	Q	R	M	D
GAG	GAG	GAG	GAT	GGA	GCG	GGC	GCC	GAG	GAG	TCG	GGA	CAG	CCC	CGT	AGC	TTC	ACG
E	E	E	D	G	A	G	A	E	E	S	G	Q	P	R	S	F	T
CAG	CTC	AAC	GAC	CTG	TCC	GGG	GCC	GGG	GGC	CGG	CAG	GGG	CCG	GGG	TCG	ACG	GAA
Q	L	N	D	L	S	G	A	G	G	R	Q	G	P	G	S	T	E
AAG	GAC	CCG	GGC	AGC	GCG	GAC	TCC	GAG	GCG	GAG	GGG	CTG	CCG	TAC	CCG	GCG	CTA
K	D	P	G	S	A	D	S	E	A	E	G	L	P	Y	P	A	L
GCC	CCG	GTG	GTT	TTC	TTC	TAC	TTG	AGC	CAG	GAC	AGC	CGC	CCG	CGG	AGC	TGG	TGT
A	P	V	V	F	F	Y	L	S	Q	D	S	R	P	R	S	W	C
CTC	CGC	ACG	GTC	TGT	AAC	CCG	TGG	TTC	GAG	CGA	GTC	AGT	ATG	CTG	GTC	ATT	CTT
L	R	T	V	C	N	P	W	F	E	R	V	S	M	L	V	I	L
CTC	AAC	TGT	GTG	ACT	CTG	GGT	ATG	TTC	AGG	CCG	TGT	GAG	GAC	ATT	GCC	TGT	GAC
L	N	C	V	T	L	G	M	F	R	P	C	E	D	I	A	C	D
TCC	CAG	CGC	TGC	CGG	ATC	CTG	CAG	GCC	TTC	GAT	GAC	TTG	ATC	TTT	GCC	TTC	TTT
S	Q	R	C	R	I	L	Q	A	F	D	D	F	I	F	A	F	F
GCT	GTG	GAA	ATG	GTG	GTG	AAG	ATG	GTG	GCC	TTG	GGC	ATC	TTT	GGG	AAG	AAA	TGT
A	V	E	M	V	V	K	M	V	A	L	G	I	F	G	K	K	C
TAC	CTG	GGA	GAC	ACT	TGG	AAC	CGG	CTT	GAC	TTT	TTC	ATT	GTC	ATT	GCA	GGG	ATG
Y	L	G	D	T	W	N	R	L	D	F	F	I	V	I	A	G	M

*Start of
Seq ID No: 1
+
Seq ID No: 2
(coding regions)*

DNASIS DNA Translation [T-INS]

CTG	GAG	TAT	TCG	CTG	GAC	CTG	CAG	AAC	GTC	AGC	TTC	TCC	GCA	GTC	AGG	ACA	GTC
L	E	Y	S	L	D	L	Q	N	V	S	F	S	A	V	R	T	V
CGT	GTG	CTG	CGA	CCG	GTC	AGG	GCC	ATT	AAC	CGG	GTG	CCC	AGC	ATG	CGC	ATT	CTC
R	V	L	R	P	L	R	A	I	N	R	V	P	S	M	R	I	L
GTC	ACA	TTA	CTG	CTG	GAC	ACC	TTG	CCT	ATG	CTG	GGC	AAC	GTC	CTG	CTG	CTC	TGT
V	T	L	L	L	D	T	L	P	M	L	G	N	V	L	L	L	C
TTC	TTC	GTC	TTT	TTC	ATC	TTT	GGC	ATC	GTG	GGC	GTC	CAG	CTG	TGG	GCA	GGA	CTG
F	F	V	F	F	I	F	G	I	V	G	V	Q	L	W	A	G	L
CTT	CGC	AAC	CGA	TGC	TTC	CTC	CCC	GAG	AAC	TTC	AGC	CTC	CCC	CTG	AGC	GTG	GAC
L	R	N	R	C	F	L	P	E	N	F	S	L	P	L	S	V	D
CTG	GAG	CCT	TAT	TAC	CAG	ACA	GAG	AAT	GAG	GAC	GAG	AGC	CCC	TTC	ATC	TGC	TCT
L	E	P	Y	Y	Q	T	E	N	E	D	E	S	P	F	I	C	S
CAG	CCT	CGG	GAG	AAT	GGC	ATG	AGA	TCC	TGC	AGG	AGT	GTG	CCC	ACA	CTG	CGT	GGG
Q	P	R	E	N	G	M	R	S	C	R	S	V	P	T	L	R	G
GAA	GGC	GGT	GGT	GGC	CCA	CCC	TGC	AGT	CTG	GAC	TAT	GAG	ACC	TAT	AAC	AGT	TCC
E	G	G	G	G	P	P	C	S	L	D	Y	E	T	Y	N	S	S
AGC	AAC	ACC	ACC	TGT	GTC	AAC	TGG	AAC	CAG	TAC	TAT	ACC	AAC	TGC	TCT	GCG	GGC
S	N	T	T	C	V	N	W	N	Q	Y	Y	T	N	C	S	A	G
GAG	CAC	AAC	CCC	TTC	AAA	GGC	GCC	ATC	AAC	TTT	GAC	AAC	ATT	GGC	TAT	GCC	TGG
E	H	N	P	F	K	G	A	I	N	F	D	N	I	G	Y	A	W
ATC	GCC	ATC	TTC	CAG	GTC	ATC	ACA	CTG	GAG	GGC	TGG	GTC	GAC	ATC	ATG	TAC	TTC
I	A	I	F	Q	V	I	T	L	E	G	W	V	D	I	M	Y	F
GTA	ATG	GAC	GCT	CAC	TCC	TTC	TAC	AAC	TTC	ATC	TAC	TTC	ATT	CTT	CTC	ATC	ATC
V	M	D	A	H	S	F	Y	N	F	I	Y	F	I	L	L	I	I
GTG	GGC	TCC	TTC	TTC	ATG	ATC	AAC	CTG	TGC	CTG	GTG	GTG	ATT	GCC	ACG	CAG	TTC
V	G	S	F	F	M	I	N	L	C	L	V	V	I	A	T	Q	F

DNASIS DNA Translation [T-INS]

1360	1369	1378	1387	1396	1405
TCC GAG ACC AAA	CAG CGG GAG AGT CAG	CTG ATG CGG GAG	CAG CGT GTA CGA TTC		
S E T K Q	R E S Q L M R E	Q R V R F			
1414	1423	1432	1441	1450	1459
CTG TCC AAT GCT AGC	ACC CTG GCA AGC TTC	TCT GAG CCA GGC AGC	TGC TAT GAG		
L S N A S T	L A S F S E P G S C Y E				
1468	1477	1486	1495	1504	1513
GAG CTA CTC AAG TAC	CTG GTG TAC ATC CTC	CGA AAA GCA GCC CGA	AGG CTG GCC		
E L L K Y L V Y I L R K A A R R L A					
1522	1531	1540	1549	1558	1567
CAG GTC TCT AGG GCT ATA	GGC GTG CGG GCT GGG	CTG CTC AGC AGC CCA	GTG GCC		
Q V S R A I G V R A G L L S S P V A					
1576	1585	1594	1603	1612	1621
CGT AGT GGG CAG GAG CCC	CAG CCC AGT GGC AGC TGC	ACT CGC TCA CAC	CGT CGT		
R S G Q E P Q P S G S C T R S H R R					
1630	1639	1648	1657	1666	1675
CTG TCT GTC CAC CAC	CTG GTC CAC CAC CAT	CAC CAC CAC CAT	CAC CAC CAC TAC	CAC	
L S V H H L V H H H H H H H H Y H					
1684	1693	1702	1711	1720	1729
CTG GGT AAT GGG ACG CTC	AGA GTT CCC CGG	GCC AGC CCA GAG	ATC CAG GAC	AGG	
L G N G T L R V P R A S P E I Q D R					
1738	1747	1756	1765	1774	1783
GAT GCC AAT GGG TCT CGC	CGG CTC ATG CTA CCA	CCA CCC TCT ACA	CCC ACT CCC		
D A N G S R R L M L P P P S T P T P					
1792	1801	1810	1819	1828	1837
TCT GGG GGC CCT CCG	AGG GGT GCG GAG TCT	GTA CAC AGC TTC	TAC CAT GCT	GAC	
S G G P P R G A E S V H S F Y H A D					
1846	1855	1864	1873	1882	1891
TGC CAC TTG GAG CCA	GTC CGT TGC CAG GCA	CCC CCT CCC AGA	TGC CCA TCG	GAG	
C H L E P V R C Q A P P P R C P S E					
1900	1909	1918	1927	1936	1945
GCA TCT GGT AGG ACT	GTG GGT AGT GGG AAG	GTG TAC CCC ACT	GTG CAT ACC	AGC	
A S G R T V G S G K V Y P T V H T S					
1954	1963	1972	1981	1990	1999
CCT CCA CCA GAG ATA	CTG AAG GAT AAA GCA	CTA GTG GAG GTG	GCC CCC AGC	CCT	
P P P E I L K D K A L V E V A P S P					
2008	2017	2026	2035	2044	2053
GGG CCC CCC ACC CTC	ACC AGC TTC AAC	ATC CCA CCT GGG	CCC TTC AGC	TCC ATG	
G P P T L T S F N I P P G P S S M					

DNASIS DNA Translation [T-INS]

2062	2071	2080	2089	2098	2107
CAC AAG CTC CTG GAG ACA CAG AGT ACG GGA GCC TGC CAT AGC TCC TGC AAA ATC					
H K L L E T Q S T G A C H S S C K I					
2116	2125	2134	2143	2152	2161
TCC AGC CCT TGC TCC AAG GCA GAC AGT GGA GCC TGC GGG CCG GAC AGT TGT CCC					
S S P C S K A D S G A C G P D S C P					
2170	2179	2188	2197	2206	2215
TAC TGT GCC CGG ACA GGA GCA GGA GAG CCA GAG TCC GCT GAC CAT GTC ATG CCT					
Y C A R T G A G E P E S A D H V M P					
2224	2233	2242	2251	2260	2269
GAC TCA GAC AGC GAG GCT GTG TAT GAG TTC ACA CAG GAC GCT CAG CAC AGT GAC					
D S D S E A V Y E F T Q D A Q H S D					
2278	2287	2296	2305	2314	2323
CTC CGG GAT CCC CAC AGC CGG CGG CGA CAG CGG AGC CTG GGC CCA GAT GCA GAG					
L R D P H S R R R Q R S L G P D A E					
2332	2341	2350	2359	2368	2377
CCT AGT TCT GTG CTG GCT TTC TGG ACG CTG ATC TGT GAC ACA TTC CGG AAG ATC					
P S S V L A F W R L I C D T F R K I					
2386	2395	2404	2413	2422	2431
GTA GAT AGC AAA TAC TTT GGC CGG GGA ATC ATG ATC GCC ATC CTG GTC AAT ACA					
V D S K Y F G R G I M I A I L V N T					
2440	2449	2458	2467	2476	2485
CTC AGC ATG GGC ATC GAG TAC CAC GAG CAG CCC GAG GAG CTC ACC AAC GCC CTG					
L S M G I E Y H E Q P E E L T N A L					
2494	2503	2512	2521	2530	2539
GAA ATC AGC AAC ATC GTC TTC ACC AGC CTC TTC GCC TTG GAG ATG CTG CTG AAA					
E I S N I V F T S L F A L E M L L K					
2548	2557	2566	2575	2584	2593
CTG CTT GTC TAC GGT CCC TTT GGC TAC ATT AAG AAT CCC TAC AAC ATC TTT GAT					
L L V Y G P F G Y I K N E Y N I F D					
2602	2611	2620	2629	2638	2647
GGT GTC ATT GTG GTC ATC AGT GTG TGG GAG ATT GTG GGC CAG CAG GGA GGT GGC					
G V I V V I S V W E I V G Q Q G G G					
2656	2665	2674	2683	2692	2701
CTG TCG GTG CTG CGG ACC TTC CGC CTG ATG CGG GTG CTG AAG CTG GTG CGC TTC					
L S V L R T F R L M R V L K L V R F					
2710	2719	2728	2737	2746	2755
CTG CCG GCC CTG CAG CGC CAG CTC GTG GTG CTC ATG AAG ACC ATG GAC AAC GTG					
L P A L Q R Q L V V L M K T M D N V					

DNASIS DNA Translation [T-INS]

2764	2773	2782	2791	2800	2809
GCC ACC TTC TGC ATG CTC	CTC ATG CTG TTC ATC TTC	ATC TTC AGC ATC CTG GGC			
A T F C M L L M L F I F I F S I L G					
2818	2827	2836	2845	2854	2863
ATG CAT CTC TTT GGT TGC	AAG TTC GCA TCT GAA CGG	GAT GGG GAC ACG TTG CCA			
M H L F G C K F A S E R D G D T L P					
2872	2881	2890	2899	2908	2917
GAC CGG AAG AAT TTC GAC	TCC CTG CTC TGG GCC ATC	GTC ACT GTC TTT CAG ATT			
D R K N F D S L L W A I V T V F Q I					
2926	2935	2944	2953	2962	2971
CTG ACT CAG GAA GAC TGG	AAT AAA GTC CTC TAC AAC	GGC ATG GCC TCC ACA TCG			
L T Q E D W N K V L Y N G M A S T S					
2980	2989	2998	3007	3016	3025
TCT TGG GCT GCT CTT TAC	TTT ATC GCC CTC ATG ACT	TTT GGC AAC TAT GTG CTC			
S W A A L Y F I A L M T F G N Y V L					
3034	3043	3052	3061	3070	3079
TTT AAC CTG CTG GTG GCC	ATT CTT GTG GAA GGA TTC	CAG GCA GAG GAA ATC GGC			
F N L L V A I L V E G F Q A E E I G					
3088	3097	3106	3115	3124	3133
AAA CGG GAA GAT GCG AGT	GGA CAG TTA AGC TGT ATT	CAG CTG CCT GTC AAC TCT			
K R E D A S G Q L S C I Q L P V N S					
3142	3151	3160	3169	3178	3187
CAG GGG GGA GAT GCC ACC	AAG TCT GAG TCA GAG CCT	GAT TTC TTT TCG CCC AGT			
Q G G D A T K S E S E P D F F S P S					
3196	3205	3214	3223	3232	3241
GTG GAT GGT GAT GGG GAC	AGA AAG AAG CGC TTG GCC	CTG GTG GCT TTG GGA GAA			
V D G D G D R K K R L A L V A L G E					
3250	3259	3268	3277	3286	3295
CAC GCG GAA CTA CGA AAG	AGC CTT TTG CCA CCC CTC	ATC ATC CAT ACG GCT GCG			
H A E L R K S L L P P L I I H T A A					
3304	3313	3322	3331	3340	3349
ACA CCA ATG TCA CTA CCC	AAG AGC TCC AGC ACA GGT	GTG GGG GAA GCA CTG GGC			
T P M S L P K S S S T G V G E A L G					
3358	3367	3376	3385	3394	3403
TCT GGC TCT CGA CGT ACC	AGT AGC AGT GGG TCC GCT	GAG CCT GGA GCT GCC CAC			
S G S R R T S S S G S A E P G A A H					
3412	3421	3430	3439	3448	3457
CAT GAG ATG AAA TCT CCG	CCA AGT GCC CGC AGC TCC	CCG CAC AGT CCC TGG AGT			
H E M K S P P S A R S S P H S P W S					

DNASIS DNA Translation [T-INS]

3466	3475	3484	3493	3502	3511
GCG GCA AGC AGC TGG ACC AGC AGG CGC TCC AGC AGG AAC AGC CTG GGC CGG GCC					
A A S S W T S R R S S R N S L G R A					
3520	3529	3538	3547	3556	3565
CCC AGC CTA AAG CGG AGG AGC CCG AGC GGG GAG CGG AGG TCC CTG CTG TCT GGA					
P S L K R R S P S G E R R S L L S G					
3574	3583	3592	3601	3610	3619
GAG GGC CAG GAG AGT CAG GAT GAG GAG GAA AGT TCA GAA GAG GAC CGG GCC AGC					
E G Q E S Q D E E E S S E E D R A S					
3628	3637	3646	3655	3664	3673
CCA GCA GGC AGT GAC CAT CGC CAC AGG GGT TCC TTG GAA CGT GAG GCC AAG AGT					
P A G S D H R H R G S L E R E A K S					
3682	3691	3700	3709	3718	3727
TCC TTT GAC CTG CCT GAC ACT CTG CAG GTG CCG GGG CTG CAC CGC ACA GCC AGC					
S F D L P D T L Q V P G L H R T A S					
3736	3745	3754	3763	3772	3781
GGC CGG AGC TCT GCC TCT GAG CAC CAA GAC TGT AAT GGC AAG TCG GCT TCA GGG					
G R S S A S E H Q D C N G K S A S G					
3790	3799	3808	3817	3826	3835
CGT TTG GCC CGC ACC CTG AGG ACT GAT GAC CCC CAA CTG GAT GGG GAT GAT GAC					
R L A R T L R T D D P Q L D G D D D					
3844	3853	3862	3871	3880	3889
AAT GAT GAG GGA AAT CTG AGC AAA GGG GAA CGC ATA CAA GCC TGG GTC AGA TCC					
N D E G N L S K G E R I Q A W V R S					
3898	3907	3916	3925	3934	3943
CGG CTT CCT GCC TGT TGC CGA GAG CGA GAT TCC TGG TCG GCC TAT ATC TTT CCT					
R L P A C C R E R D S W S A Y I F P					
3952	3961	3970	3979	3988	3997
CCT CAG TCA AGG TTT CGT CTC CTG TGT CAC CGG ATC ATC ACC CAC AAG ATG TTT					
P Q S R F R L L C H R I I T H K M F					
4006	4015	4024	4033	4042	4051
GAC CAT GTG GTC CTC GTC ATC ATC TTC CTC AAC TGT ATC ACC ATC GCT ATG GAG					
D H V V L V I I F L N C I T I A M E					
4060	4069	4078	4087	4096	4105
CGC CCC AAA ATT GAC CCC CAC AGC GCT GAG CGC ATC TTC CTG ACC CTC TCC AAC					
R P K I D P H S A E R I F L T L S N					
4114	4123	4132	4141	4150	4159
TAC ATC TTC ACG GCA GTC TTT CTA GCT GAA ATG ACA GTG AAG GTG GTG GCA CTG					
Y I F T A V F L A E M T V K V V A L					

DNASIS DNA Translation [T-INS]

4168	4177	4186	4195	4204	4213
GGC TGG TGC TTT GGC GAG CAG GCC TAC CTG CGC AGC AGC TGG AAT GTG CTG GAC					
G W C F G E Q A Y L R S S W N V L D					
4222	4231	4240	4249	4258	4267
GGC TTG CTG GTG CTC ATC TCC GTC ATC GAC ATC CTG GTC TCC ATG GTC TCC GAC					
G L L V L I S V I D I L V S M V S D					
4276	4285	4294	4303	4312	4321
AGC GGC ACC AAG ATC CTT GGC ATG CTG AGG GTG CTG CGG CTG CTG CGG ACC CTG					
S G T K I L G M L R V L R L L R T L					
4330	4339	4348	4357	4366	4375
CGT CCA CTC AGG GTC ATC AGC CGG GCC CAG GGA CTG AAG CTG GTG GTA GAG ACT					
R P L R V I S R A Q G L K L V V E T					
4384	4393	4402	4411	4420	4429
CTG ATG TCA TCC CTC AAA CCC ATT GGC AAC ATT GTG GTC ATT TGC TGT GCC TTC					
L M S S L K P I G N I V V I C C A F					
4438	4447	4456	4465	4474	4483
TTC ATC ATT TTT GGA ATT CTC GGG GTG CAG CTC TTC AAA GGG AAG TTC TTC GTG					
F I I F G I L G V Q L F K G K F F V					
4492	4501	4510	4519	4528	4537
TGT CAG GGT GAG GAC ACC AGG AAC ATC ACT AAC AAA TCC GAC TGC GCT GAG GCC					
C Q G E D T R N I T N K S D C A E A					
4546	4555	4564	4573	4582	4591
AGC TAC CGA TGG GTC CGG CAC AAG TAC AAC TTT GAC AAC CTG GGC CAG GCT CTG					
S Y R W V R H K Y N F D N L G Q A L					
4600	4609	4618	4627	4636	4645
ATG TCC CTG TTT GTG CTG GCC TCC AAG GAT GGT TGG GTT GAC ATC ATG TAT GAT					
M S L F V L A S K D G W V D I M Y D					
4654	4663	4672	4681	4690	4699
GGG CTG GAT GCT GTG GGT GTG GAT CAG CAG CCC ATC ATG AAC CAC AAC CCC TGG					
G L D A V G V D Q Q P I M N H N P W					
4708	4717	4726	4735	4744	4753
ATG CTG CTA TAC TTC ATC TCC TTC CTC CTC ATC GTG GCC TTC TTT GTC CTG AAC					
M L L Y F I S F L L I V A F F V L N					
4762	4771	4780	4789	4798	4807
ATG TTT GTG GGC GTG GTG GTG GAG AAC TTC CAT AAG TGC AGA CAG CAC CAG GAG					
M F V G V V V E N F H K C R Q H Q E					
4816	4825	4834	4843	4852	4861
GAG GAG GAG GCG AGG CGG CGT GAG GAG AAG CGA CTA CGG AGG CTG GAG AAA AAG					
E E E A R R R E E K R L R R L E K K					

DNASIS DNA Translation [T-INS]

4870	4879	4888	4897	4906	4915
AGA AGG AAT CTA	ATG TTG GAC GAT GTA ATT	GCT TCC GGC AGC TCA GCC	AGC GCT		
R R N L	M L D D V I A S	G S S A S A			
4924	4933	4942	4951	4960	4969
GCG TCA GAA GCC CAG TGC AAG CCC TAC TAC TCT GAC TAC TCG AGA TTC CGG CTC					
A S E A Q C K P Y Y S D Y S R F R L					
4978	4987	4996	5005	5014	5023
CTT GTC CAC CAC CTG TGT ACC AGC CAC TAC CTG GAC CTC TTC ATC ACT GGT GTC					
L V H H L C T S H Y L D L F I T G V					
5032	5041	5050	5059	5068	5077
ATC GGG CTG AAC GTG GTC ACT ATG GCC ATG GAA CAT TAC CAG CAG CCC CAG ATC					
I G L N V V T M A M E H Y Q Q P Q I					
5086	5095	5104	5113	5122	5131
CTG GAC GAG GCT CTG AAG ATC TGC AAT TAC ATC TTT ACC GTC ATC TTT GTC TTT					
L D E A L K I C N Y I F T V I F V F					
5140	5149	5158	5167	5176	5185
GAG TCA GTT TTC AAA CTT GTG GGC TTT GGC TTC CGC CGT TTC TTC CAG GAC AGG					
E S V F K L V A F G F R R F F Q D R					
5194	5203	5212	5221	5230	5239
TGG AAC CAG CTG GAC CTG GCT ATT CTG CTT CTG TCC ATC ATG GGC ATC ACA CTG					
W N Q L D L A I V L L S I M G I T L					
5248	5257	5266	5275	5284	5293
GAG GAG ATT GAG GTC AAT GCT TCG CTG CCC ATC AAC CCC ACC ATC ATC CGT ATC					
E E I E V N A S L P I N P T I I R I					
5302	5311	5320	5329	5338	5347
ATG AGG GTG CTC CGC ATT GCT CGA GTT CTG AAG CTG TTG AAG ATG GCT GTG GGC					
M R V L R I A R V L K L L K M A V G					
5356	5365	5374	5383	5392	5401
ATG CGG GCA CTG CTG GAC ACG GTG ATG CAG GGC CTG CCC CAG GTG GGG AAC CTG					
M R A L L D T V M Q A L P Q V G N L					
5410	5419	5428	5437	5446	5455
GGA CTT CTC TTC ATG TTA TTG TTT TTC ATC TTT GCA GCT CTG GGC GTG GAG CTC					
G L L F M L L F F I F A A L G V E L					
5464	5473	5482	5491	5500	5509
TTT GGA GAC CTG GAG TGT GAT GAG ACA CAC CCT TGT GAG GGC TTG GGT CGG CAT					
F G D L E C D E T H P C E G L G R H					
5518	5527	5536	5545	5554	5563
GCC ACC TTT AGG AAC TTT GGT ATG GCC TTT CTG ACC CTC TTC CGA GTC TCC ACT					
A T F R N F G M A F L T L F R V S T					

DNASIS DNA Translation [T-INS]

5572	5581	5590	5599	5608	5617
GGT GAC AAC TGG AAT GGT ATT ATG AAG GAC ACC CTC CGG GAC TGT GAC CAG GAG					
G D N W N G I M K D T L R D C D Q E					
5626	5635	5644	5653	5662	5671
TCC ACC TGC TAC AAC ACT GTC ATC TCC CCT ATC TAC TTT GTG TCC TTC GTG CTG					
S T C Y N T V I S P I Y F V S F V L					
5680	5689	5698	5707	5716	5725
ACG GCC CAG TTT GTG CTG GTC AAC GTG GTC ATA GCT GTG CTG ATG AAG CAC CTG					
T A Q F V L V N V V I A V L M K H L					
5734	5743	5752	5761	5770	5779
GAA GAA AGC AAC AAA GAG GCT AAG GAG GAG GCC GAG CTC GAG GCC GAG CTG GAG					
E E S N K E A K E E A E L E A E L E					
5788	5797	5806	5815	5824	5833
CTG GAG ATG AAG ACG CTC AGC CCG CAG CCC CAC TCC CCG CTG GGC AGC CCC TTC					
L E M K T L S P Q P H S P L G S P F					
5842	5851	5860	5869	5878	5887
CTC TGG CCC GGG GTG GAG GGT GTC AAC AGT CCT GAC AGC CCT AAG CCT GGG GCT					
L W P G V E G V N S P D S P K P G A					
5896	5905	5914	5923	5932	5941
CCA CAC ACC ACT GCC CAC ATT GGA GCA GCC TCG GGC TTC TCC CTT GAG CAC CCC					
P H T T A H I G A A S G F S L E H P					
5950	5959	5968	5977	5986	5995
ACG ATG GTA CCC CAC CCC GAG GAG GTG CCA GTC CCC CTA GGA CCA GAC CTG CTG					
T M V P H P E E V P V P L G P D L L					
6004	6013	6022	6031	6040	6049
ACT GTG AGG AAG TCT GGT GTC AGC CGG ACG CAC TCT CTG CCC AAT GAC AGC TAC					
T V R K S G V S R T H S L P N D S Y					
6058	6067	6076	6085	6094	6103
ATG TGC CGC AAT GGG AGC ACT GCT GAG AGA TCC CTA GGA CAC AGG GGC TGG GGG					
M C R N G S T A E R S L G H R G W G					
6112	6121	6130	6139	6148	6157
CTC CCC AAA GCC CAG TCA GGC TCC ATC TTG TCC GTT CAC TCC CAA CCA GCA GAC					
L P K A Q S G S I L S V H S Q P A D					
6166	6175	6184	6193	6202	6211
ACC AGC TGC ATC CTA CAG CTT CCC AAA GAT GTG CAC TAT CTG CTC CAG CCT CAT					
T S C I L Q L P K D V H Y L L Q P H					
6220	6229	6238	6247	6256	6265
GGG GCC CCC ACC TGG GGC GCC ATC CCT AAA CTA CCC CCA CCT GGC CGC TCC CCT					
G A P T W G A I P K L P P P G R S P					

DNASIS DNA Translation [T-INS]

6274	6283	6292	6301	6310	6319
CTG GCT CAG AGG	CCT CTC AGG CGC CAG GCA	GCA ATA AGG ACT GAC TCC	CTG GAT		
L A Q R	P L R R Q A A I R T D S L D				
6328	6337	6346	6355	6364	6373
GTG CAG GGC CTG GGT AGC CGG GAA GAC CTG TTG TCA GAG GTG AGT GGG CCC TCC					
V Q G L G S R E D L L S E V S G P S					
6382	6391	6400	6409	6418	6427
TGC CCT CTG ACC CGG TCC TCA TCC TTC TGG GGC GGG TCG AGC ATC CAG GTG CAG					
C P L T R S S S F W G G S S I Q V Q					
6436	6445	6454	6463	6472	6481
CAG CGT TCC GGC ATC CAG AGC AAA GTC TCC AAG CAC ATC CGC CTG CCA GCC CCT					
Q R S G I Q S K V S K H I R L P A P					
6490	6499	6508	6517	6526	6535
TGC CCA GGC CTG GAA CCC AGC TGG GCC AAG GAC CCT CCA GAG ACC AGA AGC AGC					
C P G L E P S W A K D P P E T R S S					
6544	6553	6562	6571	6580	6589
TTA GAG CTG GAC ACG GAG CTG AGC TGG ATT TCA GGA GAC CTC CTT CCC AGC AGC					
L E L D T E L S W I S G D L L P S S					
6598	6607	6616	6625	6634	6643
CAG GAA GAA CCC CTG TCC CCA CGG GAC CTG AAG AAG TGC TAC AGT GTA GAG ACC					
Q E E P L S P R D L K K C Y S V E T					
6652	6661	6670	6679	6688	6697
CAG AGC TGC AGG CGC AGG CCT GGG TCC TGG CTA GAT GAA CAG CGG AGA CAC TCC					
Q S C R R R P G S W L D E Q R R H S					
6706	6715	6724	6733	6742	6751
ATT GCT GTC AGC TGT CTG GAC AGC GGC TCC CAA CCC CGC CTA TGT CCA AGC CCC					
I A V S C L D S G S Q P R L C P S P					
6760	6769	6778	6787	6796	6805
TCA AGC CTC GGG GGC CAA CCT CTT GGG GGT CCT GGG AGC CGG CCT AAG AAA AAA					
S S L G G Q P L G G P G S R P K K K					
6814	6823	6832	6841	6850	6859
CTC AGC CCA CCC AGT ATC TCT ATA GAC CCC CCG GAG AGC CAG GGC TCT CGG CCC					
L S P P S I S I D P P E S Q G S R P					
6868	6877	6886	6895	6904	6913
CCA TGC AGT CCT GGT GTC TGC CTC AGG AGG AGG GCG CCG GCC AGT GAC TCT AAG					
P C S P G V C L R R R A P A S D S K					
6922	6931	6940	6949	6958	6967
GAT CCC TCG GTC TCC AGC CCC CTT GAC AGC ACG GCT GCC TCA CCC TCC CCA AAG					
D P S V S S P L D S T A A S P S P K					

DNASIS DNA Translation [T-INS]

6976	6985	6994	7003	7012	7021
AAA GAC ACG CTG AGT CTC	TCT GGT TTG TCT TCT GAC	CCA ACA GAC ATG GAC CCC			
K D T L S L S G L S S D P T D M D P					
7030	7039	7048	7057	7066	7075
TGA GTC CTA CCC ACT CTC	CCC CAT CAC CTT TCT CCA CCG	GGT GCA GAT CCT AGC			
* V L P T L P H H L S P P G A D P S					
7084	7093	7102	7111	7120	7129
TCC GCC TCC TGG GCA GCG TTT	CTG AAA AGT CCC ACG TAA	GCA GCA AGC AGC CAC			
S A S W A A F L K S P T * A A S S H					
7138	7147	7156	7165	7174	7183
GAG GCA CCT CAC CTG CCT TCT	TGA GTG GCT GGT GGG GAT	GAC GAG CAG AAC TTC			
E A P H L P S S V A G G D D E Q N F					
7192	7201	7210	7219	7228	7237
CGG AGA GTC GAT CTG AAG	AGA ACA CAG CCC TGG AGC	CCC TGC CTC CGG GAA GAA			
R R V D L K R T Q P W S P C L R E E					
7246	7255	7264	7273	7282	
GGA AAA GGA GAA AGC CCA	GTG TGG CCA AGG CTC CCG	ACA CCA GGA GCT G/3'			
G K G E S P V W P R L P T P G					

seq 1 + 3
A/seq 2 + 4

T-INS

10	20	30	40	50	60
GEAGCTGAGC	TGAACTGGCC	CTCCTGGGGA	CTCAGCAAGC	TCTCTAGAGC	CCCCCATGT
70	80	90	100	110	120
CTCCCCCACC	GGGGTCCCC	GGTTGCGTGA	GGACACCTCC	TCTGAGGGGC	TCCGCTCGCC
130	140	150	160	170	180
CCTCTTCGGA	CCCCCGGG	CCCCGGCTGG	CCAGAGGATG	GACGAGGAGG	AGGATGGAGC
190	200	210	220	230	240
GGGCGCCGAG	GAGTCGGGAC	AGCCCCGTAG	CTTCACGCAG	CTCAACGACC	TGTCGGGGGC
250	260	270	280	290	300
CGGGGGCCGG	CAGGGGCCGG	GGTCGACGGA	AAAGGACCCG	GGCAGCGCGG	ACTCCGAGGC
310	320	330	340	350	360
GGAGGGGCTG	CCGTACCCGG	CGCTAGCCCC	GGTGGTTTTC	TTCTACTTGA	GCCAGGACAG
370	380	390	400	410	420
CCGCCCCGCG	AGCTGGTGTC	TCGGCACGGT	CTGTAACCCG	TGGTTCGAGC	GAGTCAGTAT
430	440	450	460	470	480
GCTGGTCATT	CTTCTCAACT	GTGTGACTCT	GGGTATGTTC	AGGCCGTGTG	AGGACATTGC
490	500	510	520	530	540
CTGTGACTCC	CAGCGCTGCC	GGATCCTGCA	GGCCTTCGAT	GACTTCATCT	TTGCCTTCTT
550	560	570	580	590	600
TGCTGTGGAA	ATGGTGGTGA	AGATGGTGCC	CTTGGGCATC	TTTGGGAAGA	AATGTTACCT
610	620	630	640	650	660
GGGAGACACT	TGGAACCGGC	TTGACTTTTT	CATTGTCTATT	GCAGGGATGC	TGGAGTATTC
670	680	690	700	710	720
GCTGGACCTG	CAGAACGTCA	GCTTCTCCGC	AGTCAGGACA	GTCCGTGTGC	TGCGACCGCT
730	740	750	760	770	780
CAGGGCCATT	AACCGGTGTC	CCAGCATGCG	CATTCTCGTC	ACATTACTGC	TGGACACCTT
790	800	810	820	830	840
GCCTATGCTG	GGCAACGTCC	TGCTGCTCTG	TTTCTTCGTC	TTTTTCATCT	TTGGCATCGT
850	860	870	880	890	900
GGGCGTCCAG	CTGTGGGCAG	GACTGCTTCG	CAACCGATGC	TTCTCTCCCG	AGAACTTCAG
910	920	930	940	950	960
CCTCCCCCTG	AGCGTGGACC	TGGAGCCTTA	TTACCAGACA	GAGAATGAGG	ACGAGAGCCC
970	980	990	1000	1010	1020
CTTCATCTGC	TCTCAGCCTC	GGGAGAATGG	CATGAGATCG	TGCAGGAGTG	TGCCCCACT
1030	1040	1050	1060	1070	1080
GCGTGGGGAA	GGCGGTGGTG	GCCCACCCTG	CAGTCTGGAC	TATGAGACCT	ATAACAGTTC
1090	1100	1110	1120	1130	1140
CAGCAACACC	ACCTGTGTCA	ACTGGAACCA	GTACTATACC	AACTGTCTCTG	CGGGCGAGCA
1150	1160	1170	1180	1190	1200
CAACCCCTTC	AAAGCGCCA	TCAACTTTGA	CAACATTGGC	TATGCTGGA	TCGCCATCTT
1210	1220	1230	1240	1250	1260
CCAGGTATC	AACTGGAGG	GCTGGGTGCA	CATCATGTAC	TTGCTAATGG	ACGCTCACTC

- Start of
 SEQ ID NO: 1
 coding region

[illegible]

Sub
a
cont

DNASIS
T-INS

1270	1280	1290	1300	1310	1320
CTTCTACAAC	TTTCATCTACT	TCATCTCTCT	CATCATCGTG	GGCTCCTTCT	TCATGATCAA
1330	1340	1350	1360	1370	1380
CCTGTGGCTG	GTGGTGATTG	CCACGCAGTT	CTCCGAGACC	AAACAGCGGG	AGAGTCAGCT
1390	1400	1410	1420	1430	1440
GATGCGGGAG	CAGCGTGTAC	GATTCTCTGTC	CAATGCTAGC	ACCCTGGCAA	GCTTCTCTGA
1450	1460	1470	1480	1490	1500
GCCAGGCAGC	TGCTATGAGG	AGCTACTCAA	GTACCTGGTG	TACATCTCTC	GAAAAGCAGC
1510	1520	1530	1540	1550	1560
CCGAAGGCTG	GCCCAGGTCT	CTAGGGCTAT	AGGCGTGCGG	GCTGGGCTGC	TCAGCAGCCC
1570	1580	1590	1600	1610	1620
AGTGGCCCGT	AGTGGGCAGG	AGCCCCAGCC	CAGTGGCAGC	TGCACTCGCT	CACACCGTCG
1630	1640	1650	1660	1670	1680
TCTGTCTGTC	CACCACCTGG	TCCACCACCA	TCACCACCAC	CATCACCCTC	ACCACCTGGG
1690	1700	1710	1720	1730	1740
TAATGGGACG	CTCAGAGTTC	CCCGGGCCAG	CCCAGAGATC	CAGGACAGGG	ATGCCAATGG
1750	1760	1770	1780	1790	1800
GTCTCGCCGG	CTCATGCTAC	CACCACCTCT	TACACCCACT	CCCTCTGGGG	GCCCTCCGAG
1810	1820	1830	1840	1850	1860
GGGTGCGGAG	TCTGTACACA	GCTTCTACCA	TGCTGACTGC	CACTTGGAGC	CAGTCCGTTC
1870	1880	1890	1900	1910	1920
CCAGGCACCC	CCTCCAGAT	GCCCATCGGA	GGCATCTGGT	AGGACTGTGG	GTAGTGGGAA
1930	1940	1950	1960	1970	1980
GGTGTACCCC	ACTGTGCATA	CCAGCCCTCC	ACCAGAGATA	CTGAAGGATA	AAGCACTAGT
1990	2000	2010	2020	2030	2040
GGAGGTGGCC	CCCAGCCCTG	GGCCCCCACC	CCTCACCAGC	TTCAACATCC	CACCTGGGCC
2050	2060	2070	2080	2090	2100
CTTCAGCTCC	ATGCACAAGC	TCCTGGAGAC	ACAGACTACG	GGAGCCTGCC	ATAGCTCTCTG
2110	2120	2130	2140	2150	2160
CAAAATCTCC	AGCCCTTGCT	CCAAGGCAGA	CAGTGGAGCC	TGCGGGCCGG	ACAGTTGTCC
2170	2180	2190	2200	2210	2220
CTACTGTGCC	CGGACAGGAG	CAGGAGAGCC	AGAGTCCGCT	GACCATGTCA	TGCCTGACTC
2230	2240	2250	2260	2270	2280
AGACAGCGAG	GCTGTGTATG	AGTTACACAC	GGACGCTCAG	CACAGTGACC	TCCGGGATCC
2290	2300	2310	2320	2330	2340
CCACAGCCGG	CGGCGACAGC	GGAGCCTGGG	CCCAGATGCA	GAGCCTAGTT	CTGTGCTGGC
2350	2360	2370	2380	2390	2400
TTTCTGGAGG	CTGATCTGTG	ACACATTCCG	GAAGATCGTA	GATAGCAAAT	ACTTTGGCCG
2410	2420	2430	2440	2450	2460
GGGAATCATG	ATCGCCATCC	TGGTCAATAC	ACTCAGCATG	GGCATCGAGT	ACCACGAGCA
2470	2480	2490	2500	2510	2520
GCCCGAGGAG	CTCACCAACG	CCCTGGAAAT	CAGCAACATC	GTCTTCACCA	GCCTCTTCGC

DNASIS
T-INS

2530	2540	2550	2560	2570	2580
CTTGGAGATG	CTGCTGAAAC	TGCTTGCTCTA	CGGTCCCTTT	GGCTACATTA	AGAATCCCTA
2590	2600	2610	2620	2630	2640
CAACATCTTT	GATGGTGTC	TGTGGTTCAT	CAGTGTGTGG	GAGATTGTGG	GCCAGCAGGG
2650	2660	2670	2680	2690	2700
AGGTGGCCTG	TGGTGCTGC	GGACCTTCCG	CCTGATGCGG	GTGCTGAAGC	TGGTGCGCTT
2710	2720	2730	2740	2750	2760
CCTGCCGGCC	CTGCAGCGCC	AGCTCGTGGT	GCTCATGAAG	ACCATGGACA	ACGTGGCCAC
2770	2780	2790	2800	2810	2820
CTTCTGCATG	CTCCTCATGC	TGTTTCATCTT	CATCTTCAGC	ATCCTGGGCA	TGCATCTCTT
2830	2840	2850	2860	2870	2880
TGTTTGCAAG	TTCGCATCTG	AACGGGATGG	GGACACGTTG	CCAGACCGGA	AGAATTTCGA
2890	2900	2910	2920	2930	2940
CTCCCTGCTC	TGGGCCATCG	TCACTGTCTT	TCAGATTCTG	ACTCAGGAAG	ACTGGAATAA
2950	2960	2970	2980	2990	3000
AGTCCTCTAC	AACGGCATGG	CCTCCACATC	GTCTTGGGCT	GCTCTTTACT	TCATCGCCCT
3010	3020	3030	3040	3050	3060
CATGACTTTT	GGCAACTATG	TGCTCTTTAA	CCTGCTGGTG	GCCATTCTTG	TGGAAGGATT
3070	3080	3090	3100	3110	3120
CCAGGCAGAG	GAAATCGGCA	AACGGGAAGA	TGCGAGTGG	CAGTTAAGCT	GTATTACAGCT
3130	3140	3150	3160	3170	3180
GCCTGTCAAC	TCTCAGGGGG	GAGATGCCAC	CAAGTCTGAG	TCAGAGCCTG	ATTTCTTTTC
3190	3200	3210	3220	3230	3240
GCCCAGTGTG	GATGGTGATG	GGGACAGAAA	GAAGCGCTTG	GCCCTGGTGG	CTTTGGGAGA
3250	3260	3270	3280	3290	3300
ACACGCGGAA	CTACGAAAGA	GCCTTTTGCC	ACCCCTCATC	ATCCATACGG	CTGCGACACC
3310	3320	3330	3340	3350	3360
AATGTCACTA	CCCAAGAGCT	CCAGCACAGG	TGTGGGGGAA	GCACTGGGCT	CTGGCTCTCG
3370	3380	3390	3400	3410	3420
ACGTACCAGT	AGCAGTGGGT	CCGCTGAGCC	TGGAGCTGCC	CACCATGAGA	TGAAATCTCC
3430	3440	3450	3460	3470	3480
GCCAAGTGCC	CGCAGCTCCC	CGCACAGTCC	CTGGAGTGGG	GCAAGCAGCT	GGACCAGCAG
3490	3500	3510	3520	3530	3540
GCGCTCCAGC	AGGAACAGCC	TGGGCCGGGC	CCCCAGCCTA	AAGCGGAGGA	GCCCAGCGGG
3550	3560	3570	3580	3590	3600
GGAGCGGAGG	TCCCTGCTGT	CTGGAGAGGG	CCAGGAGAGT	CAGGATGAGG	AGGAAAGTTC
3610	3620	3630	3640	3650	3660
AGAAGAGGAC	CGGGCCAGCC	CAGCAGGCAG	TGACCATCGC	CACAGGGGTT	CCTTGGAACG
3670	3680	3690	3700	3710	3720
TGAGGCCAAG	AGTTCTCTTG	ACCTGCTCTA	CACCTCTGAG	GTGCCGGGGC	TGCACCGCAC
3730	3740	3750	3760	3770	3780
AGCCAGCGGC	CGGAGCTCTG	CCTCTGAGCA	CCAAGACTGT	AATGGCAAGT	CGGCTTCAGG

時間	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100
1	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100

Sub
at
cont.

DNASIS
T-INS

3790	3800	3810	3820	3830	3840
GCGTTTGGCC	CGCACCTGA	GGACTGATGA	CCCCAACTG	GATGGGGATG	ATGACAATGA
3850	3860	3870	3880	3890	3900
TGAGGGAAAT	CTGAGCAAAG	GGGAACGCAT	ACAAGCCTGG	GTCAGATCCC	GGCTTCCTGC
3910	3920	3930	3940	3950	3960
CTGTTGCCGA	GAGCGAGATT	CCTGGTCGGC	CTATATCTTT	CCTCCTCAGT	CAAGGTTTCG
3970	3980	3990	4000	4010	4020
TCTCCTGTGT	CACCGGATCA	TCACCCACAA	GATGTTTGAC	CATGTGGTCC	TCGTCAATCA
4030	4040	4050	4060	4070	4080
CTTCCTCAAC	TGTATCACCA	TCGCTATGGA	GCGCCCCAAA	ATTGACCCCC	ACAGCGCTGA
4090	4100	4110	4120	4130	4140
GCGCATCTTC	CTGACCTCT	CCAACATCAT	CTTCACGGCA	GTCTTTCTAG	CTGAAATGAC
4150	4160	4170	4180	4190	4200
AGTGAAGGTG	GTGGCACTGG	GCTGGTGCTT	TGGGGAGCAG	GCCTACCTGC	GCAGCAGCTG
4210	4220	4230	4240	4250	4260
GAATGTGCTG	GACGGCTTGC	TGGTGCTCAT	CTCCGTCATC	GACATCCTGG	TCTCCATGGT
4270	4280	4290	4300	4310	4320
CTCCGACAGC	GGCACCAAGA	TCCTTGGCAT	GCTGAGGGTG	CTGCGGCTGC	TGCGGACCCT
4330	4340	4350	4360	4370	4380
GCGTCCACTC	AGGGTCATCA	GCCGGGCCCA	GGGACTGAAG	CTGGTGGTAG	AGACTCTGAT
4390	4400	4410	4420	4430	4440
GTCATCCCTC	AAACCCATTG	GCAACATTGT	GGTCATTTCG	TGTGCCTTCT	TCATCATTTT
4450	4460	4470	4480	4490	4500
TGGAATTCTC	GGGGTGACGC	TCTTCAAAGG	GAAGTCTTTC	GTGTGTGAGG	GTGAGGACAC
4510	4520	4530	4540	4550	4560
CAGGAACATC	ACTAACAAAT	CCGACTGCGC	TGAGGCCAGC	TACCGATGGG	TCCGGCACAA
4570	4580	4590	4600	4610	4620
GTACAACTTT	GACAACCTGG	GCCAGGCTCT	GATGTCCCTG	TTTGTGCTGG	CCTCCAAGGA
4630	4640	4650	4660	4670	4680
TGGTTGGGTT	GACATCATGT	ATGATGGGCT	GGATGCTGTG	GGTGTGGATC	AGCAGCCCAT
4690	4700	4710	4720	4730	4740
CATGAACCAC	AACCCCTGGA	TGCTGCTATA	CTTCATCTCC	TACCTCTCTA	TCGTGGCCTT
4750	4760	4770	4780	4790	4800
CTTTGTCTTG	AACATGTTTG	TGGGCGTGGT	GGTGGAGAAC	TTCCATAAGT	GCAGACAGCA
4810	4820	4830	4840	4850	4860
CCAGGAGGAG	GAGGAGGCCA	GGCGGCGTGA	GGAGAAGCGA	CTACGGAGGC	TGGAGAAAAA
4870	4880	4890	4900	4910	4920
GAGAAGGAAT	CTAATGTTGG	ACGATGTAAT	TGCTTCGGGC	AGCTCAGCCA	GCGCTGCGTC
4930	4940	4950	4960	4970	4980
AGAAGCCAG	TGCAAGCCCT	ACTACTCTGA	CTACTCGAGA	TTCCGGCTCC	TTGTCCACCA
4990	5000	5010	5020	5030	5040
CCTGTGTACC	AGCCACTACC	TGGACCTCTT	CATCACTGGT	GTCATCGGGC	TGAACGTGGG

Year	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100
1980	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100

Sub
A
cont.

DNASIS
T-INS

5050	5060	5070	5080	5090	5100
CACTATGGCC	ATGGAACATT	ACCAGCAGCC	CCAGATCCTG	GACGAGGCTC	TGAAGATCTG
5110	5120	5130	5140	5150	5160
CAATTACATC	TTTACCGTCA	TCCTTGCTCT	TGAGTCAGTT	TTCAAACCTG	TGGCCTTTGG
5170	5180	5190	5200	5210	5220
CTCCGCCGT	TTCTTCCAGG	ACAGGTGGAA	CCAGCTGGAC	CTGGCTATTG	TGCTTCTGTC
5230	5240	5250	5260	5270	5280
CATCATGGGC	ATCACACTGG	AGGAGATTGA	GGTCAATGCT	TCGCTGCCCC	TCAACCCAC
5290	5300	5310	5320	5330	5340
CATCATCGCT	ATCATGAGGG	TGCTCCGCAT	TGCTCGAGTT	CTGAAGCTGT	TGAAGATGGC
5350	5360	5370	5380	5390	5400
TGTGGGCATG	CGGGCACTGC	TGGACACGGT	GATGCAGGCC	CTGCCCCAGG	TGGGGAACCT
5410	5420	5430	5440	5450	5460
GGGACTTCTC	TTCATGTTAT	TGTTTTTCAT	CTTTGCAGCT	CTGGGCGTGG	AGCTCTTTGG
5470	5480	5490	5500	5510	5520
AGACCTGGAG	TGTGATGAGA	CACACCTTGG	TGAGGGCTTG	GGTCGGCATG	CCACCTTTAG
5530	5540	5550	5560	5570	5580
GAACCTTGGT	ATGGCCTTTC	TGACCCCTCT	CCGAGTCTCC	ACTGGTGACA	ACTGGAATGG
5590	5600	5610	5620	5630	5640
TATTATGAAG	GACACCCTCC	GGGACTCTGA	CCAGGAGTCC	ACCTGCTACA	ACACTGTGAT
5650	5660	5670	5680	5690	5700
CTCCCCATC	TACTTTGTGT	CCTTCGTGCT	GACGGCCAG	TTTGTGCTGG	TCAACGTGGT
5710	5720	5730	5740	5750	5760
CATAGCTGTG	CTGATGAAGC	ACCTGGAAGA	AAGCAACAAA	GAGGCCAAGG	AGGAGGCCGA
5770	5780	5790	5800	5810	5820
GCTCGAGGCC	GAGCTGGAGC	TGGAGATGAA	GACGCTCAGC	CCGCAGCCCC	ACTCCCCGCT
5830	5840	5850	5860	5870	5880
GGGCAGCCCC	TTCTCTGGC	CCGGGGTGA	GGGTGTCAAC	AGTCCTGACA	GCCCTAAGCC
5890	5900	5910	5920	5930	5940
TGGGGCTCCA	CACACCACTG	CCCACATTGG	AGCAGCTCTG	GGCTTCTCCC	TTGAGCACCC
5950	5960	5970	5980	5990	6000
CACGATGGTA	CCCCACCCCG	AGGAGGTGCC	AGTCCCCCTA	GGACCAGACC	TGCTGACTGT
6010	6020	6030	6040	6050	6060
GAGGAAGTCT	GGTGTAGGCC	GGACGCATC	TCTGCCCAAT	GACAGCTACA	TGTGCCGCAA
6070	6080	6090	6100	6110	6120
TGGGAGCACT	GCTGAGAGAT	CCCTAGGACA	CAGGGGCTGG	GGGCTCCCCA	AAGCCAGTC
6130	6140	6150	6160	6170	6180
AGGCTCCATC	TTGTCCGTTT	ACTCCCAACC	AGCAGACACC	AGCTGCATCC	TACAGCTTCC
6190	6200	6210	6220	6230	6240
CAAAGATGTG	CACTATCTGC	TCCAGCCTCA	TGGGGCCCCC	ACCTGGGGCG	CCATCCCTAA
6250	6260	6270	6280	6290	6300
ACTACCCCCA	CCTGGCCGCT	CCCCTCTGGC	TCAGAGGCCT	CTCAGGCGCC	AGGCAGCAAT

DNASIS
T-INS

6310	6320	6330	6340	6350	6360
AAGGACTGAC	TCCCTGGATG	TGCAGGGCCT	GGGTAGCCGG	GAAGACCTGT	TGTCAGAGGT
6370	6380	6390	6400	6410	6420
GAGTGGGCCC	TCCTGCCCTC	TGACCCGGTC	CTCATCCTTC	TGGGGCGGGT	CGAGCATCCA
6430	6440	6450	6460	6470	6480
GGTGCAGCAG	CGTTCCCGCA	TCCAGAGCAA	AGTCTCCAAG	CACATCCGCC	TGCCAGCCCC
6490	6500	6510	6520	6530	6540
TTGCCCAGGC	CTGGAACCCA	GCTGGGCCAA	GGACCCTCCA	GAGACCAGAA	GCAGCTTAGA
6550	6560	6570	6580	6590	6600
GCTGGACACG	GAGCTGAGCT	GGATTTTCAGG	AGACCTCCTT	CCCAGCAGCC	AGGAAGAACC
6610	6620	6630	6640	6650	6660
CCTGTCCCCA	CGGGACCTGA	AGAAGTGCTA	CAGTGTAGAG	ACCCAGAGCT	GCAGGCGCAG
6670	6680	6690	6700	6710	6720
GCCTGGGTCC	TGGCTAGATG	AACAGCGGAG	ACACTCCATT	GCTGTCAGCT	GTCTGGACAG
6730	6740	6750	6760	6770	6780
CGGCTCCCAA	CCCCGCCTAT	GTCCAAGCCC	CTCAAGCCTC	GGGGGCCAAC	CTCTTGGGGG
6790	6800	6810	6820	6830	6840
TCCTGGGAGC	CGGCCTAAGA	AAAAACTCAG	CCCACCCAGT	ATCTCTATAG	ACCCCCCGGA
6850	6860	6870	6880	6890	6900
GAGCCAGGGC	TCTCGGCCCC	CATGCAGTCC	TGGTGTCTGC	CTCAGGAGGA	GGGCGCCGGC
6910	6920	6930	6940	6950	6960
CAGTGACTCT	AAGGATCCCT	CGGTCTCCAG	CCCCCTTGAC	AGCAGGGCTG	CCTCACCCCTC
6970	6980	6990	7000	7010	7020
CCCAAAGAAA	GACACGCTGA	GTCTCTCTGG	TTTGTCTTCT	GACCCAACAG	ACATGGACCC
7030	7040	7050	7060	7070	7080
CTGAGTCCTA	CCCACTCTCC	CCCATCACCT	TTCTCCACCG	GGTGCAGATC	CTAGCTCCGC
7090	7100	7110	7120	7130	7140
CTCCTGGGCA	GCGTTTCTGA	AAAGTCCCAC	GTAAGCAGCA	AGCAGCCACG	AGGCACCTCA
7150	7160	7170	7180	7190	7200
CCTGCCTTCT	TCACTGGCTG	GTGGGGATGA	CGACGAGAAC	TTCCGGAGAG	TCGATCTGAA
7210	7220	7230	7240	7250	7260
GAGAACACAG	CCCTGGAGCC	CCTGCCTCCG	GGAAGAAGGA	AAAGGAGAAA	GCCCAGTGTG
7270	7280	7290	7300	7310	7320
GCCAAGGCTC	CCGACACCAG	GAGCTG

end of
seq 1+3

start of SEQ ID NO:2 (coding region)

ELS*TGPPGDSASSLEPPPTCSPTGVPRRLREDTSSSEGLRSLFLFGPPG
 APAGQRDEEEDGAGAEESGQPRSFTQLNDLSGAGGRQGPGSTEK
 DPGSADSEAEGLPYPALAPVVFYLSQDSRPRWSCLRTVCNPWFE
 RVSMLVILLNVTLTGMFRPCEDIACDSQRCRILQAFDDFIAFFAV
 EMVVMKVALGIFGKKCYLGDWTNRLLDFFIVIAMGLEYSLDLQNV
 FSAVRTVRVLRPLRAINRVPSMRILVTLTLLDTLPMLGNVLLCFFV
 FFIIGIVGVQLWAGLLRNRCFLPENFSLPLSVDLEPYQYQTENEDES
 PFICSQPRENGMRSCRSVPTLRGEGGGGPPCSLDYETYNSSSNTT
 CVNWNQYTYNCSAGEHNPFKGAINFDNIGYAWIAIFQVITLEGWV
 DIMYFVMDAHSFYNFYIFILLIIVGSFFMINCLCVVIATQFSETKQR
 ESQLMREQVRVRLSNASTLASFSEPGSCYEELLKYL VYILRKAAR
 RLAQVSRAGVRAQLLSSPVARSQEPQPSGSCSTRSHRRLSVHHL
 VH HHHHHHHHHYHLNGTLRVPRASPEIQDRDANGSRRLMLPPPST
 TPTSGGPPRGAESEVHSFYHADCHLEPVRQCAPPPRCPSEASGRTV
 GSGKVYPTVHTSTPPPEILKDKALVEVAPSGPPTLTSFNIPPGPFS
 MHKLLTQSTGACHSSCKISSPCSKADSGACGPDSCPYCARTGAG
 EPESADHVMPDSDEAVYEFTQDAQHSDLRDPHSRRRQSRSLGPD
 EPSSVLAFWRLICDTRFKIVDSKYFGRGIMIAILVNTLSMGIEYHEQ
 PEELTNALEISNIVFTSLFALEMLLKLLVYGPFYIKNPYNIFDGV
 VVISVWEIVGQGGGLSVLRTFRLMRVLKLVRLPALQRLVVLVLM
 KTMDNVATFCMLLMFLFIFISILGMHLFGCKFASERDGD TLPDRK
 NFD SLLWAIVTVFQILTQEDWNKVLYNGMASTSSWAALYFIALMT
 FGNYVLFNLLVAILVEGFQAEIEIGKREDASGQLSCIQLPVNSQGGD
 ATKSESEPDFFSVSDGDGRKKRLALVALGEHAELRKSLLPLII
 HTAATPMSLPKSSSTGVGEALGSGSRRTSSSGSAEPGAHHHEMKS
 PPSARSSPHSPWSAASSWTSRRSSRNSLGRAPSLKRRSPSGERRS
 LLSGEGQESQDEEESSEEDRASPA GSDHRHRGSLEREAKSSFDLPD
 TLQVPLGHRTASGRSSASEHQCNGKSASGRLARTLRTDDPQLDG
 DDDNDEGNLSKGERIQAWVRSLR PACCERDSWSAYIFPPQSRFR
 LLCHRIITHKMFHDHVVVLVIIFLNCITIAMERP KIDPHSAERIFLTLN
 YIFTAVFLAEMTVKVVALGWCFGEQAYLRSSWNVLDGLLVLSVI
 DILYSMVSDSGTKILGMLRVLRLRLTLRPLRVISRAQGLKLVVETL
 MSSLKPIGNIVVICCAFFIIFGILGVQLFKGKFFVCQGEDTRNITNK
 SDCAEASRYVRHKYNFDNLGLQALMSLFVLASKDGWVDIMYDGL
 DAVGVDDQPPIMNHNPNWMLLYFISELLIVAFFVLNMFVGVVVENFH
 KCRQHQEEEEEARREEKRRLRLEKKRRNLMLDDVIASGSSASAAS
 EAQCKPYSDYSRFRLLVHHLCTSHYLDL FITGVIGLNVVTMAME
 HYQQPQILDEALKICNYIFTVIFVFSVFKLVAFGFRFFQDRWNQ
 LDLAIVLLSIMGITLEEIEVNASLPINPTIIRIMRVLRIRAVLKLKLM
 AVGMRALLDTVMQALPQVGNLGLLFLMLLFFIFAALGVELFGDLEC
 DETHPCEGLGRHATFRNFGMAFLTFRVSTGDNWNGIMKDTLRDC
 DQUESTCYNTVISPIYFVSFVLTAQFVLNVVIAVLMKHLEESNKEA
 KEEAELEAELEMKTLSPQPHSLGSPFLWPGVEGVNSPDSPKPG
 APHTTAHIGAASGSFSEHPTMVPHPEEVVPVLPGLD LTVRKSGVSR
 THSLPNDSYMCNRNGSTAERSLGHRGWGLPKAQSGSILSVHSQPAD
 TSCILQLPKDVHYLLQPHGAPTWGAIPKLPPPGRSPLAQRP LRRQA
 AIRTDSLDDVQGLGSRD LLEVSVPSCPTRSSFWGGSSIQVQQR
 SGIQSKVSKHIRLPAPCPGLEPSWAKDPPESTRSSLELDTLSWISG
 DLLPSSQEEPLSPRDLKKCYSVETQSCRRRPGSWLDEQRRHSIAV
 SCLDSGSQPRLCPSPSLLGGQPLGGPGSRKKKKLSPPSISIDPPESQ
 GSRPPCSPGVCLRRRAPASDSKDPVS SPLDSTAASPSPKKDTLSL
 SGLSSDPTDMDP*VLPTLPHHLSPPGADPSSASWAAFLKSPT* AAS
 SH207IEAPHLPSYVAGGDDEQNFRFVLDKRTQPWSPCLREEGKGE
 SPVWPRLPTPGA/

SEQ 2 + 4

SEQ ID NO:5: ~~SKEKQMA~~

SEQ ID NO:6: 5' TNGC(A/C/T)ATGGAG(C/A)GNCC(C/T)-3'

SEQ ID NO:7: 5'-CTT(C/G/T)CCCTTGAA(G/C)A(G/A)CTG)-3'

SEQ ID NO:8: 5'-CCGCTGTCGGAGACCATGGAGACC-3'

SEQ ID NO:9: 5'-AGCGGCCCAAATTGACCCCCACAG-3'

SEQ ID NO:10: 5'-GAAGATGCGAGTGGACAG-3'

SEQ ID NO:11: 5'- CTGTGCGGATGGTCACTG-3'

Sub
cc
ante

[illegible]